



# NASA's GeneLab Phase II: Federated Search and Data Discovery

**Daniel C. Berrios MD PhD MPH**  
**Sylvain Costes PhD**  
**Peter Tran MS**

**NASA Ames Research Center**





# Agenda



- GeneLab
- Federated Search
  - Common Metadata Model
  - Metadata Export
- Next Steps



- Goals
  - An integrated repository and bioinformatics data system for analysis and modeling
  - Enable the discovery and validation of molecular networks that are influenced by space conditions through ground-based and flight research using next-generation omics technologies
  - Engage the broadest possible community of researchers, industry, and the general public to foster innovation
  - Strengthen international partnerships by leveraging existing capabilities and data sharing



# Phased Implementation



## Phase 1 Searchable Data FY2014 –2015

### Data System

- ✓ Public Website
- ✓ Searchable Data Repository
- ✓ Top Level Requirements
- ✓ New Data and Legacy Data

## Phase 2 Data Exchange FY2016-2017

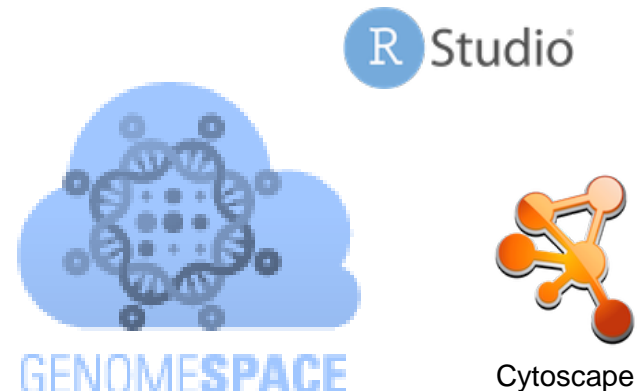
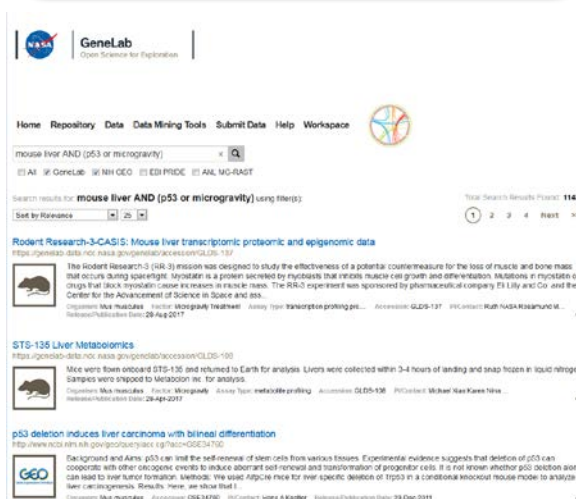
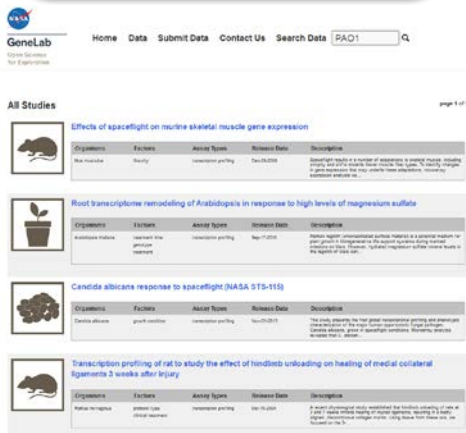
### Data System

- **Link to Public Databases via Metadata Federation**
- **Integrated Search**

## Phase 3 System Integration FY2018– 2019

### Data System

- Provide collaboration framework and tools
- Build Community via collaborative science analysis & modeling
- Provide access to biocomputational tools for omics analysis







- What?
  - Search using 1 system over multiple data sources
  - For example, Google web search
- Why?
  - Facilitates discovery of data similar to known data
  - Improves search efficiency: no need to switch and search multiple source systems
- How?
  - Metadata Mapping of Data Sources
  - If systems have search interfaces:
    - Dynamic query translation
  - If systems do not have search interfaces, or for greater reliability:
    - Metadata warehousing



Accession No	SSE60473
Title	Gravitational signature of synchronized cell cultures in particular cell cycle stages
Summary	Cell cycle and cell proliferation are decoupled under altered gravity conditions...
Organism	Arabidopsis thaliana
Experiment Type	Expression profiling by array
Status	Public on Aug 23, 2014



Accession No	SSE60473
Study Title	Gravitational signature of synchronized cell cultures in particular cell cycle stages
Study Description	Cell cycle and cell proliferation are decoupled under altered gravity conditions...
Organism	Arabidopsis thaliana
Assay Type	Transcription Profiling by Array
Release Date	23-Aug-2014



Accession No	PXD005035
Title	Mouse muscle LC-MS/MS upon weightlessness
Description	Upon weightlessness and microgravity, deleterious effects on the neurosensory and neurovestibular systems, haematological changes...
Species	Mus musculus (Mouse)
Tissue	muscle
Experiment Type	Shotgun proteomics



Accession No	mgp4026
principle investigator	Peter Girguis, Harvard University
Description	-
funding source	NASA
Visibility	Public



# Common Metadata Model



Accession	GLDS-131
Study Title	Rodent Research-3-CASIS: <b>Mouse</b> liver transcriptomic proteomic and epigenomic data
Study Description	The Rodent Research-3 (RR-3) mission was designed to study the effectiveness of a potential countermeasure for the loss of muscle and bone mass that occurs during spaceflight. <b>Myostatin</b> is a protein secreted by myoblasts that inhibits muscle cell growth and differentiation. Mutations in myostatin or drugs that block <b>myostatin</b> cause increases in muscle mass. The RR-3 experiment was sponsored by pharmaceutical company Eli Lilly and Co. and the Center for the Advancement of Science in Space and assessed the efficacy of <b>myostatin</b> inhibition to prevent skeletal muscle atrophy and weakness....



Accession	GSE466
Title	mRNA expression in regenerated mdx <b>mouse</b> skeletal muscle
Summary	... A fourfold decrease in <b>myostatin</b> mRNA in the mdx muscle was noted. Differential upregulation of actin-related protein 2/3 (subunit 4), beta-thymosin, calponin, mast cell chymase, and guanidinoacetate methyltransferase mRNA in the more benign mdx was also observed. ...



“mouse”  
“myostatin”



Home Repository Data Data Mining Tools Submit Data Help Workspace



mouse myostatin x Q

☐ All ☒ GeneLab ☒ NIH GEO ☐ EBI PRIDE ☐ ANL MG-RAST

Search results for: **mouse myostatin** using filter(s):

Total Search Results Found: 7

Sort by Relevance 25

1

## Myostatin inactivation effects on myogenesis in vitro and in vivo

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28986>



Key words: dystrophin, mdx mouse, Duchenne, fibrosis, dystrophy ABSTRACT Stimulating the commitment of implanted dystrophin+ muscle derived stem cells (MDSC) into myogenic, as opposed to lipofibrogenic, lineages is a promising therapeutic strategy for Duchenne muscular dystrophy (DMD). To examine whether counteracting myostatin, a negative regulator of muscle mass and a pro-lipofibrotic factor, would help this process, we compared the in vitro myogenic and fibrogenic capacity of MDSC from wild...

Organism: Mus musculus Accession: GSE28986 PI/Contact: Robert Gelfand Release/Publication Date: 30-Sep-2012

## The transcriptomic signature of myostatin inhibitory influence on the differentiation of mouse C2C12 myoblasts

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE59674>



GDF8 (myostatin) is a unique cytokine strongly affecting the skeletal muscle phenotype in human and animals. The aim of the present study was to elucidate the molecular mechanism of myostatin influence on the differentiation of mouse C2C12 myoblasts, using the global-transcriptome analysis with the DNA microarray technique. Treatment with exogenous GDF8 strongly affected the growth and development of C2C12 mouse myoblasts. This was manifested by the inhibition of proliferation and differentiation...

Organism: Mus musculus Accession: GSE59674 PI/Contact: Zofia Wikik Release/Publication Date: 23-Jul-2014

## Development of gene expression signature for defining the cell potency of muscle derived stem cells (MDSC) from mice of different genotypes

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE39765>



In order to determine the cell potency, by identification of genes responsible for pluri/multi potency, we performed a global gene expression profiling of MDSC isolated from five week old male wild type(WT), C57B6J and another hypertrophied musculature mouse genotype called myostatin null (Mstn<sup>-/-</sup>) mice using microarray analysis and compared this gene expression to that of a standard mouse ES cell line W4. Muscle derived stem cells (MDSC) were isolated from WT and Mstn null mice using an esta...

Organism: Mus musculus Accession: GSE39765 PI/Contact: Bipasha Bose Release/Publication Date: 01-Aug-2012

## Rodent Research-3-CASIS: Mouse liver transcriptomic proteomic and epigenomic data

<https://genelab-data.ndc.nasa.gov/genelab/accession/GLDS-137>



The Rodent Research-3 (RR-3) mission was designed to study the effectiveness of a potential countermeasure for the loss of muscle and bone mass that occurs during spaceflight. Myostatin is a protein secreted by myoblasts that inhibits muscle cell growth and differentiation. Mutations in myostatin or drugs that block myostatin cause increases in muscle mass. The RR-3 experiment was sponsored by pharmaceutical company Eli Lilly and Co. and the Center for the Advancement of Science in Space and ass...

Organism: Mus musculus Factor: Microgravity Treatment Assay Type: transcription profiling p... Accession: GLDS-137







GeneLab

Open Science for Exploration

# Federated Search Example 2



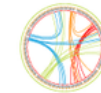
“mouse”  
“liver”



GeneLab

Open Science for Exploration

[Home](#) [Repository](#) [Data](#) [Data Mining Tools](#) [Submit Data](#) [Help](#) [Workspace](#)



mouse liver X

☐ All ☒ GeneLab ☐ NIH GEO ☒ EBI PRIDE ☐ ANL MG-RAST

Search results for: **mouse liver** using filter(s):

Total Search Results Found: **68**

Sort by Relevance 25

1 2 3 Next >>

## Rodent Research-3-CASIS: Mouse liver transcriptomic proteomic and epigenomic data

<https://genelab-data.ndc.nasa.gov/genelab/accession/GLDS-137>



The Rodent Research-3 (RR-3) mission was designed to study the effectiveness of a potential countermeasure for the loss of muscle and bone mass that occurs during spaceflight. Myostatin is a protein secreted by myoblasts that inhibits muscle cell growth and differentiation. Mutations in myostatin or drugs that block myostatin cause increases in muscle mass. The RR-3 experiment was sponsored by pharmaceutical company Eli Lilly and Co. and the Center for the Advancement of Science in Space and ass...

Organism: *Mus musculus* Factor: Microgravity Treatment Assay Type: transcription profiling p... Accession: GLDS-137  
PI/Contact: Ruth NASA Rosamund ... Release/Publication Date: 28-Aug-2017

## STS-135 Liver Metabolomics

<https://genelab-data.ndc.nasa.gov/genelab/accession/GLDS-108>



Mice were flown onboard STS-135 and returned to Earth for analysis. Livers were collected within 3-4 hours of landing and snap frozen in liquid nitrogen. Samples were shipped to Metabolon Inc. for analysis.

Organism: *Mus musculus* Factor: Microgravity Assay Type: metabolite profiling Accession: GLDS-108 PI/Contact: Michael Xiao Karen Ni...  
Release/Publication Date: 28-Apr-2017

## Female mouse liver quantitative analysis

<http://www.ebi.ac.uk/pride/archive/projects/PXD000296>



Proteomic analysis of regulated proteins in female mouse liver extracts derived from 1 week and 8 weeks of hURI expression.

Organism: *Mus musculus* (Mouse) Accession: PXD000296 PI/Contact: Pilar Ximenez-Embun Release/Publication Date: 05-Jul-2016

## The comparative proteomics research of Toxoplasma infection in mice liver

<http://www.ebi.ac.uk/pride/archive/projects/PXD003399>



The comparative proteomics research of Toxoplasma infection in mice liver

Organism: *Mus musculus* (Mouse) Accession: PXD003399 PI/Contact: Junjun He Release/Publication Date: 31-Mar-2016



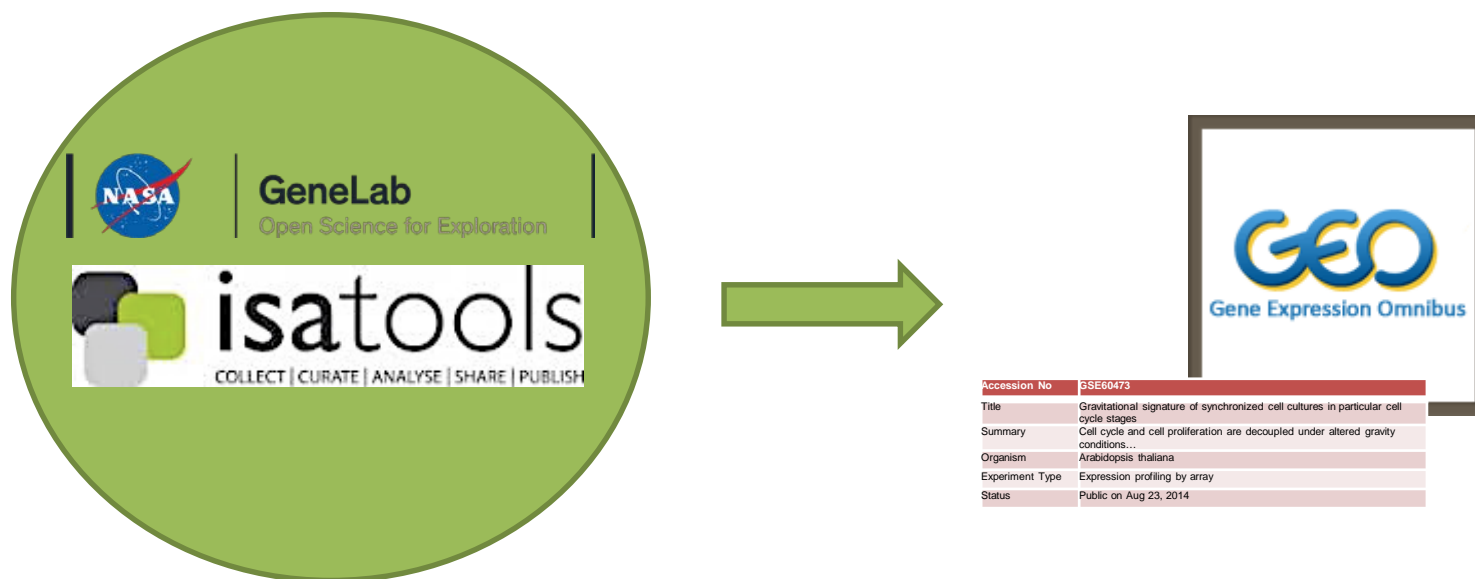
GeneLab

Open Science  
for Exploration





- Support development of federated searches initiated using extramural systems
  - Export metadata (and data, if necessary) from GeneLab to these systems
  - Provide link to “authoritative” source data (GeneLab)
- Semi-automated (scripted) process
  - GeneLab metadata used as input
  - Data products for submission to extramural data system are the output





NCBI Resources How To Sign in to NCBI

GEO DataSets GEO DataSets M-CSF space flight Search

Create alert Advanced Help

Entry type  
DataSets (0)  
Series (2)  
Samples (4)  
Platforms (1)

Organism  
Customize ...

Study type  
Expression profiling by array  
Methylation profiling by array  
Customize ...

Author  
Customize ...

Attribute name  
tissue (0)  
strain (5)  
Customize ...

Publication dates  
30 days  
1 year  
Custom range...

Clear all

Show additional filters

Summary 20 per page Sort by Default order

Search results  
Items: 7

1. [Evaluation of in vitro macrophage differentiation during space flight](#)  
(Submitter supplied) We differentiated mouse bone marrow cells in the presence of recombinant macrophage colony stimulating (rM-CSF) factor for 14 days during the flight of space shuttle Space Transportation System (STS)-126. We tested the hypothesis that the receptor expression for M-CSF, c-Fms was reduced. We used flow cytometry to assess molecules on cells that were preserved during flight to define the differentiation state of the developing bone marrow macrophages; including CD11b, CD31, CD44, Ly6C, Ly6G, F4/80, Mac2, c-Fos as well as c-Fms. [more...](#)  
Organism: Mus musculus  
Type: Expression profiling by array  
Platform: GPL1261 4 Samples  
Download data: CEL, CHP  
Series Accession: GSE95544 ID: 200095544  
[PubMed](#) [Full text in PMC](#) [Similar studies](#) [Analyze with GEO2R](#)

2. [Microarray Profile of Gene Expression during Osteoclast Differentiation in Modeled Microgravity](#)  
(Submitter supplied) Microgravity leads to a 10-15% loss of bone mass in astronauts during **space flight**. Osteoclast is the multinucleated bone resorbing cell. In this study, we used NASA developed ground based Rotary Wall Vessel Bioreactor (RWV), Rotary Cell Culture System (RCCS) to simulate microgravity ( $\mu$ Xg) conditions and demonstrated a significant increase (2-fold) in osteoclastogenesis compared to ground based control (Xg) mouse bone marrow cultures. [more...](#)  
Organism: Mus musculus  
Type: Expression profiling by array  
Platform: GPL7202 2 Samples

Filters: [Manage Filters](#)

Send to: Top Organisms [Tree](#)

Mus musculus (6)  
Homo sapiens (1)

Find related data  
Database: Select  
Find items

Search details  
("macrophage colony-stimulating factor"[MeSH Terms] OR M-CSF[All Fields]) AND ("space flight"[MeSH Terms] OR space  
Search See more...

Recent activity  
[Turn Off](#) [Clear](#)

M-CSF space flight (7) GEO DataSets  
nasa (138) GEO DataSets  
BioProject Links for SRA (Select 629115) (1) BioProject  
indoor metagenome BioProject

NCBI GEO Gene Expression Omnibus

HOME SEARCH SITE MAP GEO Publications

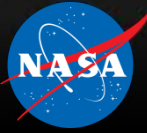
NCBI > GEO > [Accession Display](#) ?

GEO help: Mouse over screen elements for information.

Scope: Self Format: HTML Amount: Quick GEO accession: GSE95544

Series GSE95544 Query DataSets for

Status	Public on Mar 15, 2017
Title	Evaluation of in vitro macrophage differentiation during space flight
Organism	<a href="#">Mus musculus</a>
Experiment type	Expression profiling by array
Summary	We differentiated mouse bone marrow cells in the presence of recombinant macrophage colony stimulating (rM-CSF) factor for 14 days during the flight of space shuttle Space Transportation System (STS)-126. We tested the hypothesis that the receptor expression for M-CSF, c-Fms was reduced. We used flow cytometry to assess molecules on cells that were preserved during flight to define the differentiation state of the developing bone marrow macrophages; including CD11b, CD31, CD44, Ly6C, Ly6G, F4/80, Mac2, c-Fos as well as c-Fms. In addition, RNA was preserved during the flight and was used to perform a microarray. We found that there were significant differences in the expression of genes on cells that were preserved during flight compared to controls maintained on the ground. We found that there were significant changes in the distribution of genes expressed CD11b, CD31, F4/80, Mac2, Ly6C and c-Fos. However, there were no changes in c-Fms expression and no consistent pattern of advanced cell differentiation during space flight. We also found a pattern of transcript expression that would be consistent with a relatively normal differentiation outcome but proliferation by the bone marrow macrophages that were assayed after space flight. There also was a surprising pattern of space flight influence on the coagulation pathway. These data confirm that a space flight can impact on the in vitro development of macrophages from mouse bone marrow cells. MESH:Space Flight/Space Flight
Overall design	transcription profiling of 2 total treatment groups and 4 total samples
Web link	<a href="https://genelab-data.ndc.nasa.gov/genelab/accession/GLDS-50/">https://genelab-data.ndc.nasa.gov/genelab/accession/GLDS-50/</a>
Contributor(s)	Chapes SK
Citation(s)	Ortega MT, Lu N, Chapes SK. Evaluation of in vitro macrophage differentiation during space flight. <i>Adv Space Res</i> 2012 May 15;49(10):1441-1455. PMID: 23420085



- Support federated queries initiated using PRIDE, MG-RAST
  - Export metadata to these data systems
- Implement federated searches to other sources
  - MODs
  - NGOs, OGOs
- Expand search capabilities using ontologies (beyond UMLS translations) to increase discovery further





GeneLab

Open Science for Exploration

# Acknowledgements



Chris Barreras  
Afshin Beheshti  
Valery Boyko  
Sonja Caldwell  
Jairon Camarillo  
Kaushik Chakravarty  
Egle Cekanaviciute  
John Costa  
Sylvain Costes  
Marie Dinh  
Sandy Dueck  
Homer Fogel  
Jon Galazaka  
Samrawit Gebre  
Dennis Heher  
Lynn Hutchison  
Yared Kidane  
San-Huei Lai Polo  
Tristan Le  
Qiang Li  
Shu-Chun Lin  
Debora Pletcher  
Sneha Raghunandan  
Shayoni Ray  
Sigrid Reinsch  
Mike Skidmore  
David Smith  
Marla Smithwick  
Olga Stotzky

Hao Thai  
Serena Trieu  
Andrew Williamson

